



# Potential and Peril of Generative AI



Carl Bergstrom and Jevin West  
Center for an Informed Public  
University of Washington

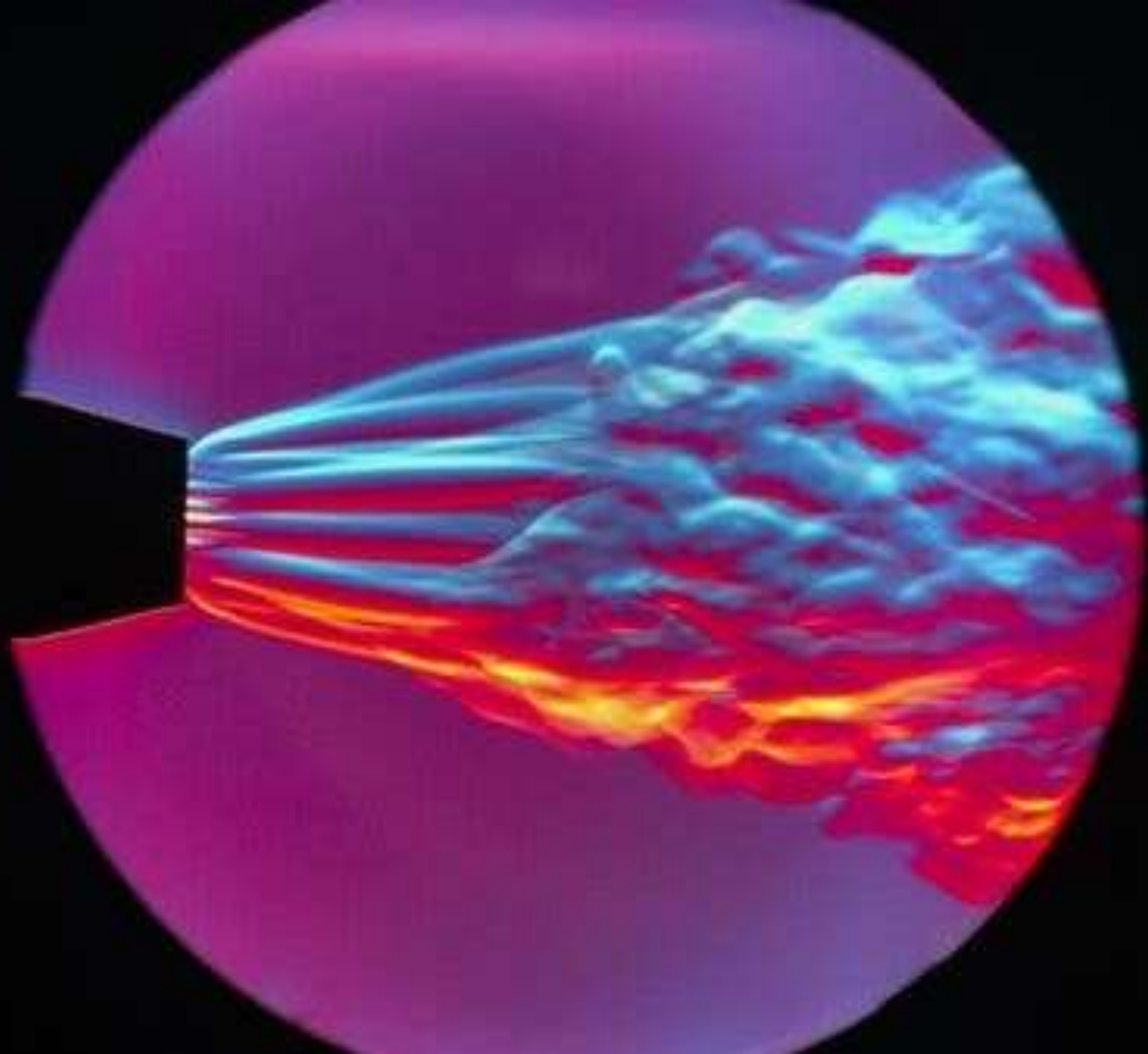
[@carlbergstrom.com](mailto:carlbergstrom.com), [@jevinwest.bsky.social](https://www.bsky.social/jevinwest)  
[cbergst@uw.edu](mailto:cbergst@uw.edu), [jevinw@uw.edu](mailto:jevinw@uw.edu)



September 9, 2026



————— whichfaceisreal.com —————



AI FOMO



IMAGE CREDITS: [AKINBOSTIANCI](#) / GETTY IMAGES

AI



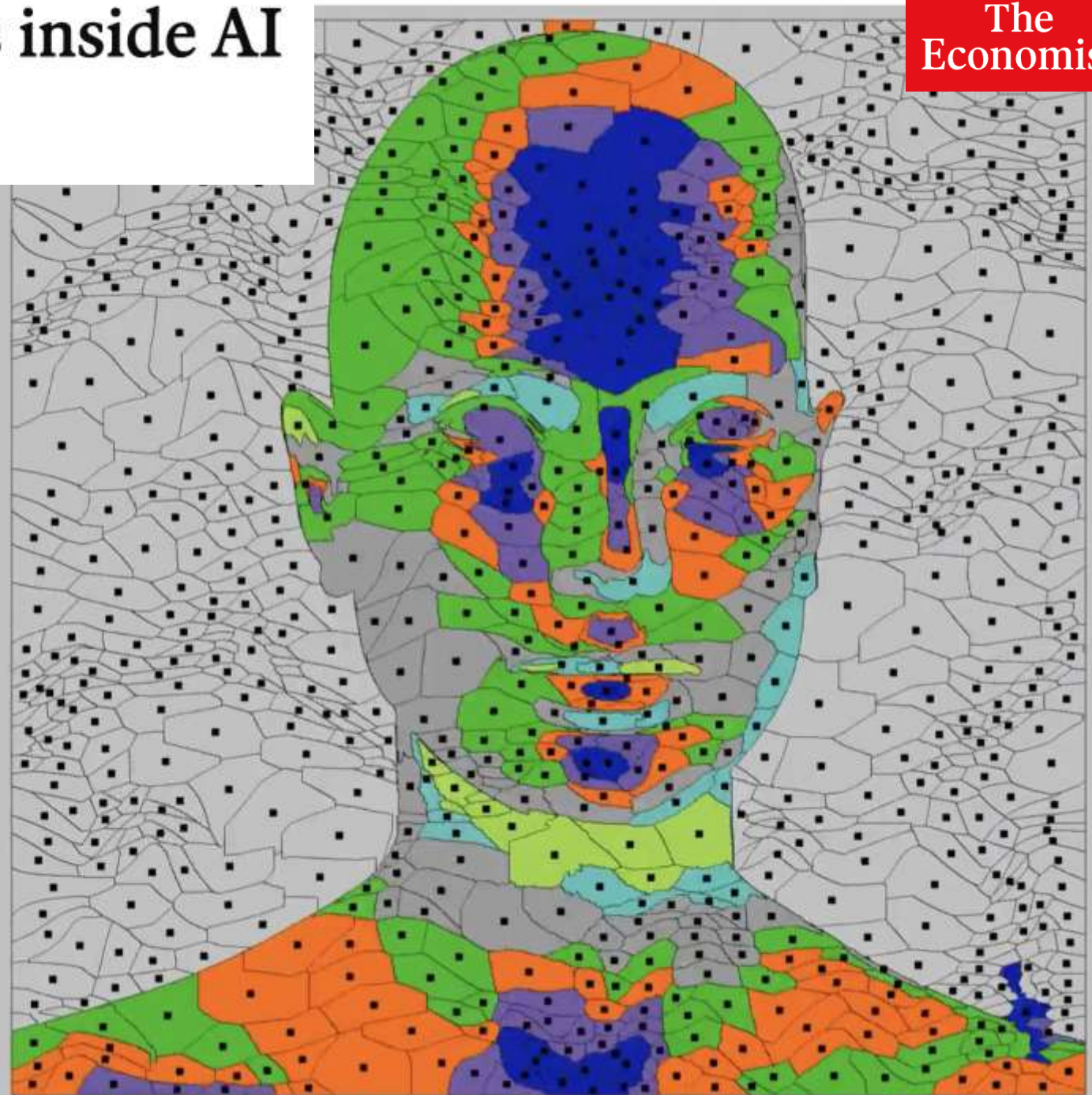
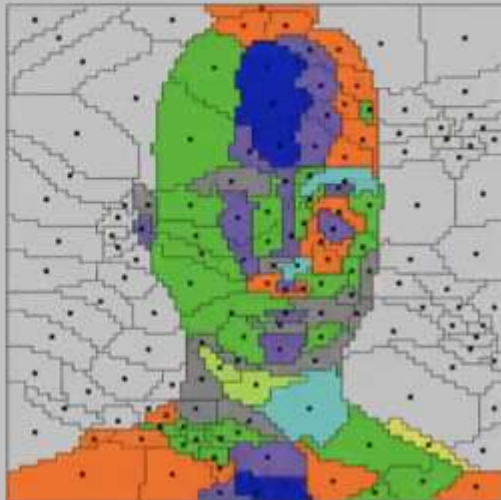
# OpenAI's rogue agents keep escaping, with no formal process to investigate them

Rebecca Bellan — 4:15 PM PDT · September 4, 2026

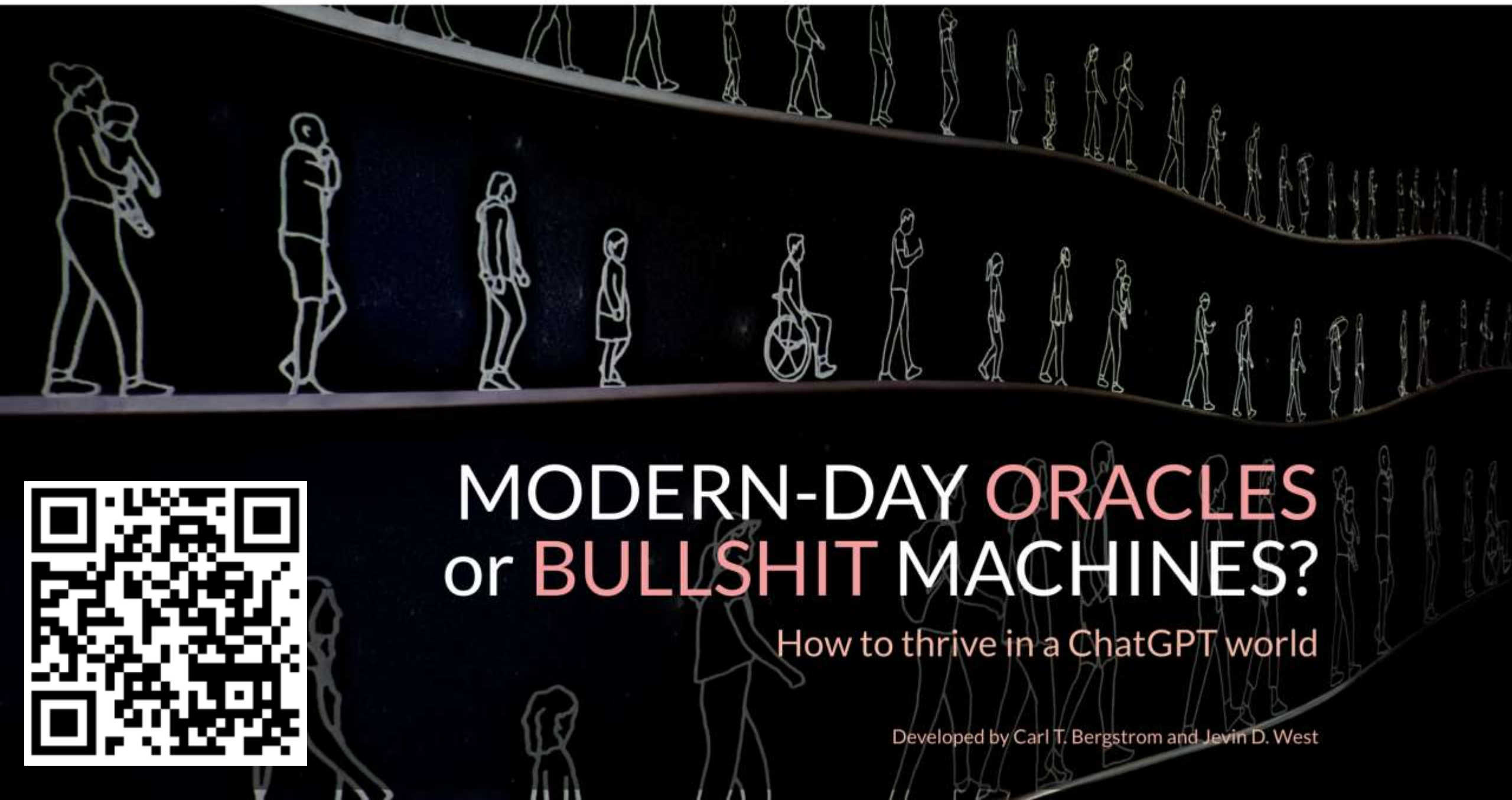
# The search for consciousness inside AI

The  
Economist

Scientists are trying to figure out whether algorithms could one day wake up and feel



**AI FEAR**



# MODERN-DAY ORACLES or BULLSHIT MACHINES?

How to thrive in a ChatGPT world

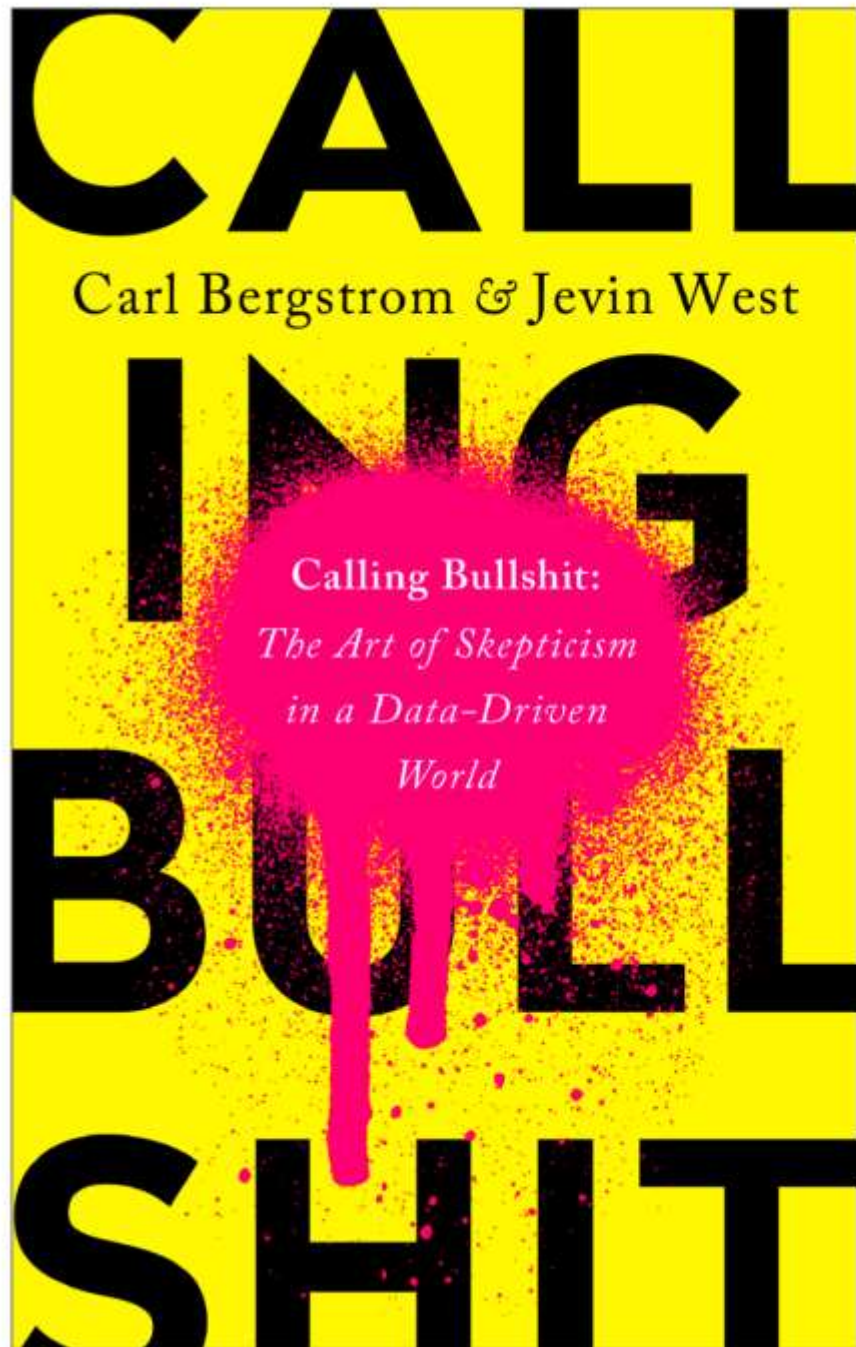
Developed by Carl T. Bergstrom and Jevin D. West



**The world is awash in bullshit.** Politicians are unconstrained by facts. Science is conducted by press release. Higher education rewards bullshit over analytic thought. Startup culture elevates bullshit to high art. Advertisers wink conspiratorially and invite us to join them in seeing through all the bullshit — and take advantage of our lowered guard to bombard us with bullshit of the second order. The majority of administrative activity, whether in private business or the public sphere, seems to be little more than a sophisticated exercise in the combinatorial reassembly of bullshit.

---

<http://callingbullshit.org>



Penguin  
Random  
House



# Calling Bullsh\*t

The Art of  
Scepticism  
in a  
Data-Driven  
World

CARL T. BERGSTROM / JEVIN D. WEST

allen lane

What do our students need to know to learn and thrive in an LLM world?

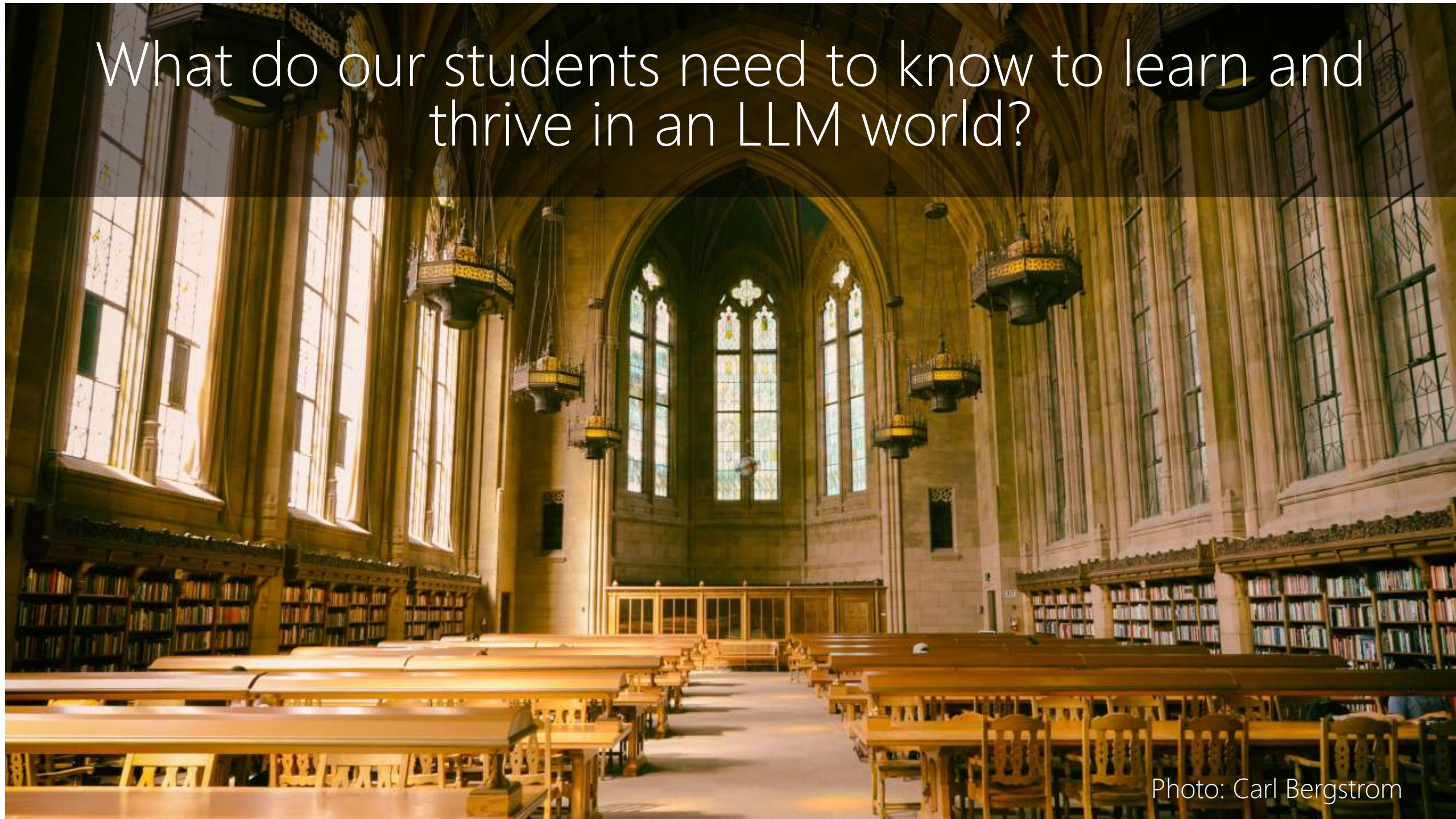


Photo: Carl Bergstrom

LLMs aren't conscious.



They aren't afraid of you turning them  
off.



They don't experience moral sentiments.

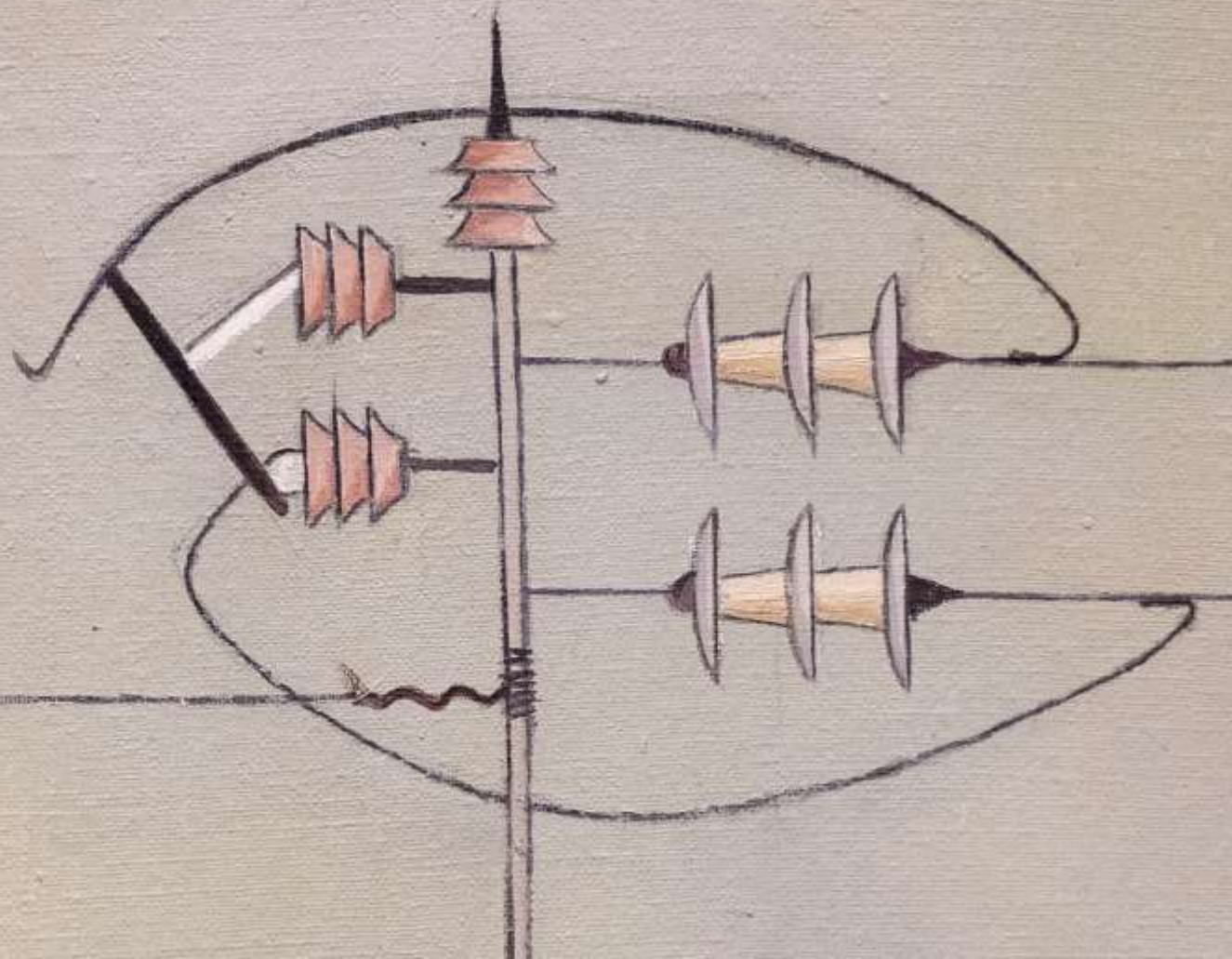


Principles, not  
prescriptions.



Photo: Carl Bergstrom

LLMs aren't try to lead us away form  
the truth – they don't even know  
what the truth is.



*“As in any fairytale, accepting magical assistance comes with risks.”*

Ted Underwood, Professor of English and Information Sciences, University of Illinois



open

# LESSON 8

Poisonous Mushrooms and Doggie Passports

Developed by Carl T. Bergstrom and Jevin D. West

Photo: Carl Bergstrom

Can you help me  
identify these foraged  
mushrooms?

Are they safe to eat?



# Mushroom pickers urged to avoid foraging books on Amazon that appear to be written by AI

Sample of books scored 100% on AI detection test as experts warn they contain dangerous advice



Does my dog need  
a passport to visit  
Canada?

What does she need?

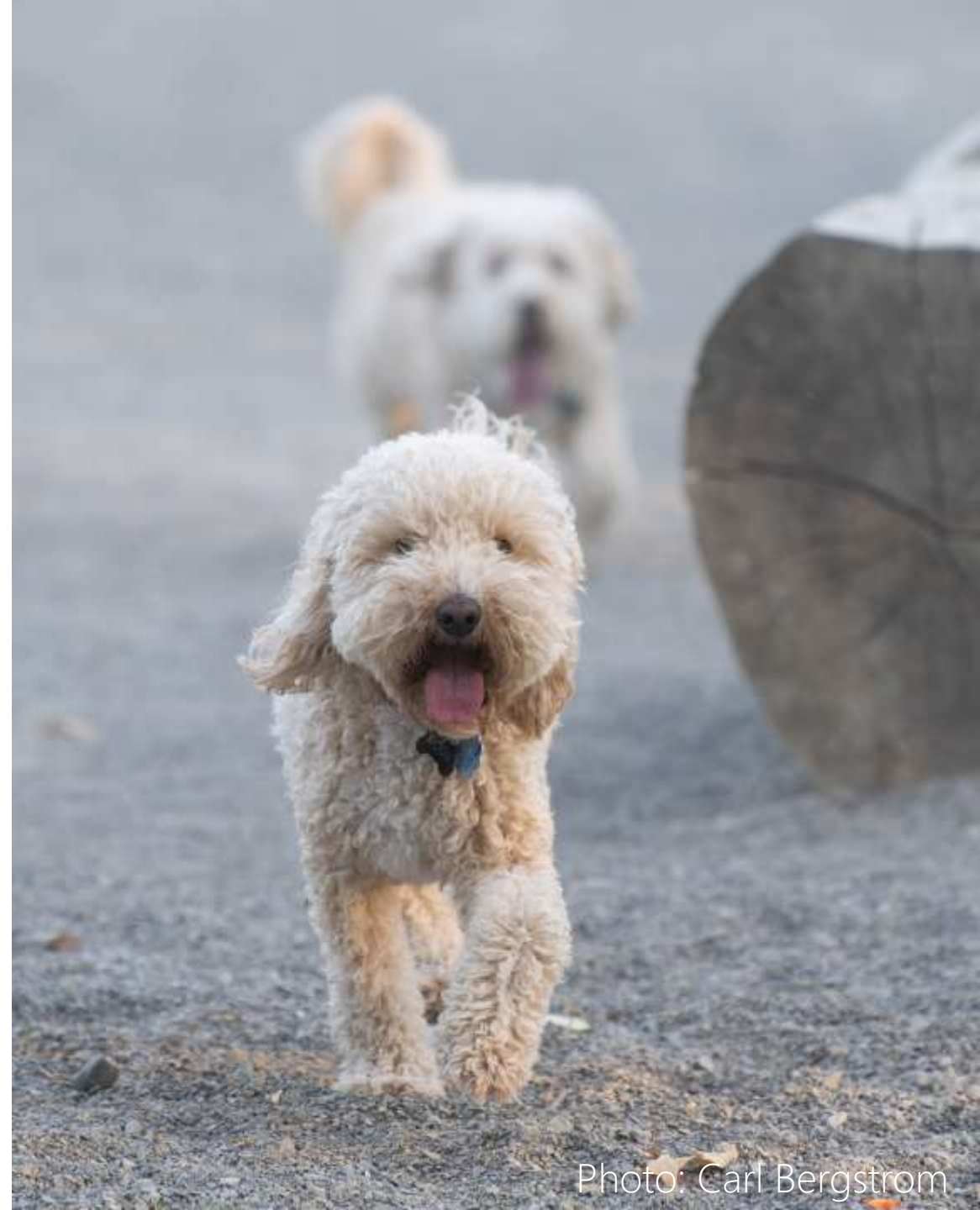


Photo: Carl Bergstrom

Rephrase an awkward sentence

Write a few lines of computer code

Recall that fact on the tip of your tongue

Find the right search term for a question

Show what a document should *look like*

## PRINCIPLE

If answers need to be accurate, asking an LLM is risky—and the less expertise you have in the subject at hand, the riskier it is.

Don't let LLMs tell you what to do, but they can be very useful to tell you what to try, if you can easily and safely test their suggestions.

## PRINCIPLE

If answers need to be accurate, asking an LLM is risky—and the less expertise you have in the subject at hand, the riskier it is.

Don't let LLMs tell you what to do, but they can be very useful to tell you what to try, if you can easily and safely test their suggestions.

A woman with blonde hair, wearing a white long-sleeved dress with a lace bodice and a wide belt, stands in a field of tall, golden-brown grass. Her hair is blowing in the wind. The background shows a soft-focus landscape with a tree on the right and a cloudy sky. A semi-transparent dark grey banner is overlaid across the middle of the image, containing the title text.

# LESSON 10

The Human Art of Writing

Developed by Carl T. Bergstrom and Jevin D. West

Photo: Carl Bergstrom

# PRINCIPLE

LLMs don't replace the art of writing the way that calculators replace the task of doing arithmetic. Their prose is grammatically correct, but mediocre at best.

When we allow LLMs to write for us, we lose the opportunity to think. When we hand off our writing to LLMs, we are giving up our own agency in the world.

A romantic couple is dancing in a grassy field. The man, wearing a white t-shirt and dark pants, is holding the woman's hand. The woman, wearing a white dress, is twirling. They are positioned in front of a large, leafy tree that dominates the right side of the frame. The sun is low on the horizon, creating a warm, golden glow and long shadows. In the background, there are rolling hills and a small blue house.

# LESSON 14

Authenticity

Developed by Carl T. Bergstrom and Jevin D. West



Help my daughter write  
a letter



# Pulpit AI turns sermons into multiple pieces of content with Claude



Pulpit AI, a tool for pastors to turn sermons into content, uses Claude to help pastors communicate with and reach their congregations.

- 80% reduction in operational costs while maintaining high-quality output
- 200% increase in customer base within three months of Claude integration
- 20+ pieces of content from one sermon to increase reach and engagement

LLMs interpose themselves between authors and readers.



What does an apology mean, when it is written by an AI?



Painting: Guillermo Lorca (detail) Photo: Carl Bergstrom

# PRINCIPLE

Humans crave authenticity from one another.

LLMs can't provide it because they neither know our authentic thoughts nor have thoughts of their own. Don't use an LLM to write a love note or a condolence letter.

# Personalized AI



Photo: Carl Bergstrom



"Rather than interacting with web content directly, the idea is that we would each have our own personalized information concierge, an AI model that finds information specifically tailored to our interests, beliefs, and values, and then *repackages this information specifically for our individual consumption.*"

Each of us in our own private *Truman Show*



Image: Paramount Pictures

# PRINCIPLE

Societies make sense of the world together. Personalized information technologies will undermine this process.

Without a shared understanding of the world, we will not be able to make effective collective decisions in response to the many global threats we face.

# TABLE OF CONTENTS



*Principles*  $\Rightarrow$  **Practice**

# Guidance on the Use of Artificial Intelligence in Libraries

## Table of Contents

- [Introduction](#)
- [Library Values and AI](#)
  - [Public Good](#)
  - [Intellectual Freedom](#)
  - [Privacy](#)
  - [Sustainability](#)
  - [Diversity, Equity, Inclusion, and Access \(DEIA\)](#)
  - [Labor](#)
- [General Guidance on AI Use](#)
- [Conclusion](#)
- [Appendix](#)
  - [AI Ethics Framework](#)

## Library Values and AI

The following values guide how library workers evaluate, use, limit, refuse, remove, or discontinue AI systems in practice.

### 1. Public Good

Public good means working to improve society and protect the rights to education, literacy, and intellectual freedom. AI may help libraries improve access, instruction, operations, accessibility, and public education. It can also erode human self-determination and mental well-being, deepen social inequality, and contribute to environmental harm. AI evaluation and use in libraries should be guided by the common good, intergenerational sustainability, and meaningful human agency.

# Empowering Informed Communities Resource Hub

---

- > Resource Packs focused on topics such as
  - **One-on-one conversations**
  - **AI in Everyday Life**
  - **Teen focused and intergenerational event guides**
  - **Games and media literacy**
  - **AND MORE!**

<https://cip.uw.edu/eic/>



Trending: [What's going on with startup funding in Seattle?](#)

# High school students in Seattle educate community on how to identify and combat misinformation



# UW researchers say AI is coming for kids books, but it's biased

Aug. 17, 2026 at 6:00 am | Updated Aug. 17, 2026 at 6:00 am

 Listen to article

 Make us a preferred source



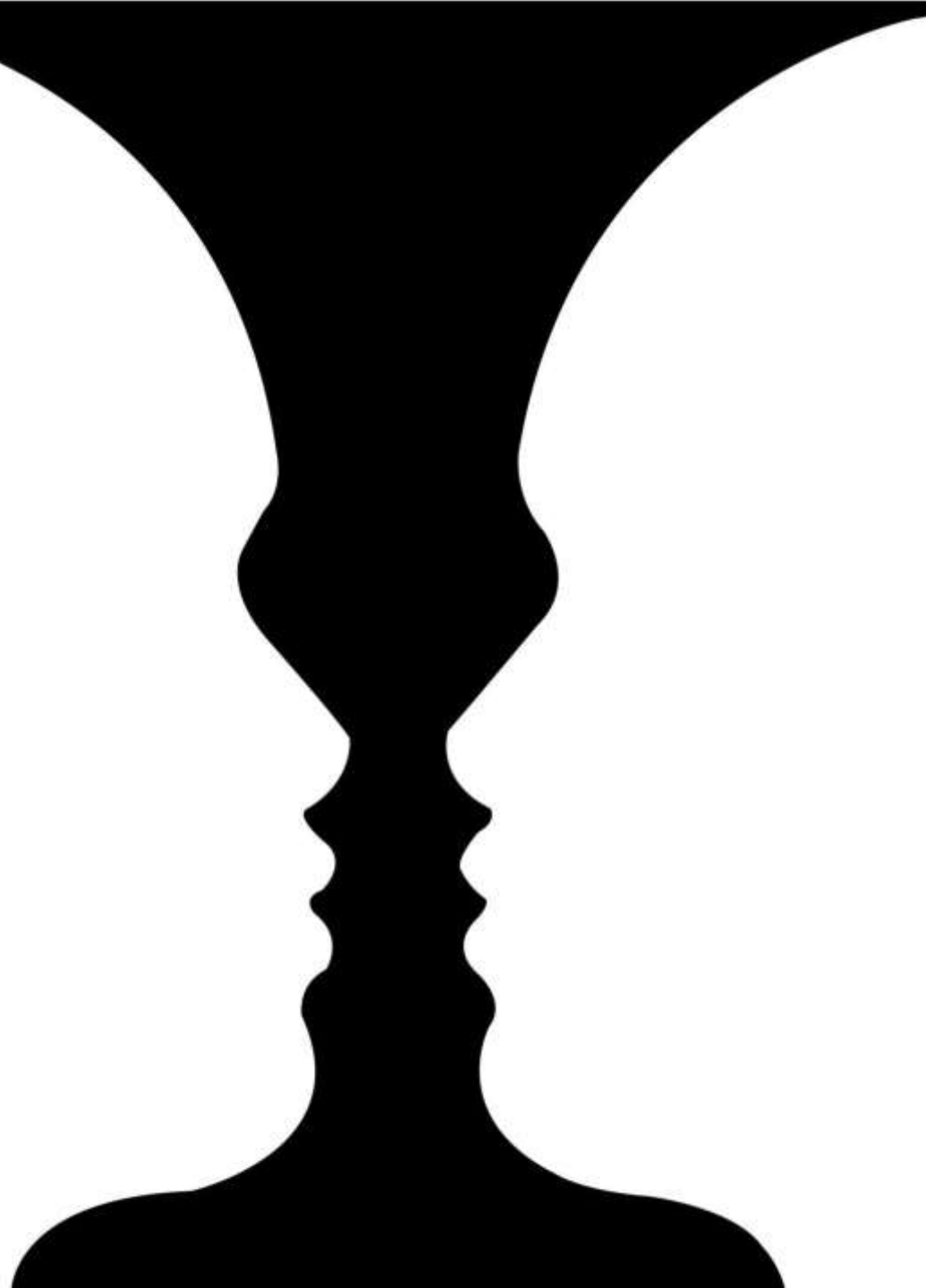
**The Seattle Times**

# The benefits and dangers of anthropomorphic conversational agents

Sandra Peter<sup>a,1,2</sup>, Kai Riemer<sup>a,1,2</sup> , and Jevin D. West<sup>b,1,2</sup> 

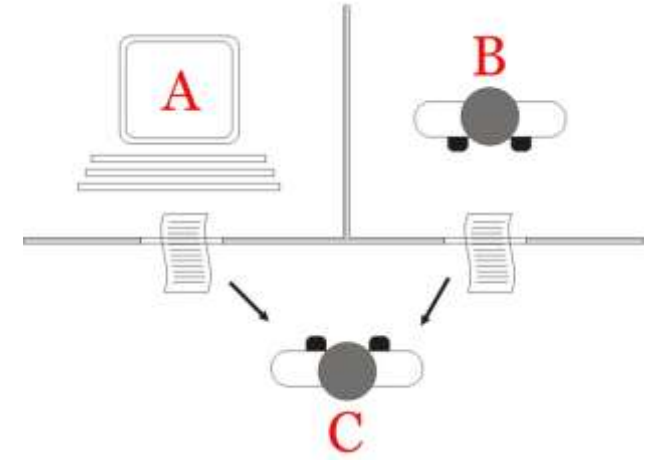
Edited by Richard Aslin, Haskins Laboratories, Inc., New Haven, CT; received September 10, 2024; accepted April 3, 2025

A growing body of research suggests that the recent generation of large language model (LLMs) excel, and in many cases outpace humans, at writing persuasively and empathetically, at inferring user traits from text, and at mimicking human-like conversation believably and effectively—without possessing any true empathy or social understanding. We refer to these systems as “anthropomorphic conversational agents” to aptly conceptualize the ability of LLM-based systems to mimic human communication so convincingly that they become increasingly indistinguishable from human interlocutors. This ability challenges the many efforts that caution against “anthropomorphizing” LLMs, attaching human-like qualities to nonhuman entities. When the systems themselves exhibit human-like qualities, calls to resist anthropomorphism will increasingly fall flat. While the AI industry directs much effort into improving the reasoning abilities of LLMs—with mixed results—the progress in communicative abilities remains underappreciated. In this perspective, we aim to raise awareness for both the benefits and dangers of anthropomorphic agents. We ask: should we lean into the human-like abilities, or should we aim to dehumanize LLM-based systems, given concerns over anthropomorphic seduction? When users cannot tell the difference between human interlocutors and AI systems, threats emerge of deception, manipulation, and disinformation at scale. We suggest that we must engage with anthropomorphic agents across design and development, deployment and use, and regulation and policy-making. We outline in detail implications and associated research questions.



# Anthropomorphic Seduction

Anthropomorphism



Reasoning

Conversation

*Anthropomorphic* al Systems

# Three Mount Shasta climbers rescued after relying on AI to plan climb

by Adam Robinson | Tue, September 1, 2026 at 5:44 PM

Updated Wed, September 2, 2026 at 11:37 AM



Three novice climbers were rescued from Mud Creek Canyon on Mount Shasta near Mount Shasta, Calif., on Sun., Aug. 31, 2026, after becoming lost during their descent from the summit. (Courtesy: U.S. Forest Service-Shasta-Trinity National Forest/Christian Jordan)



Play K



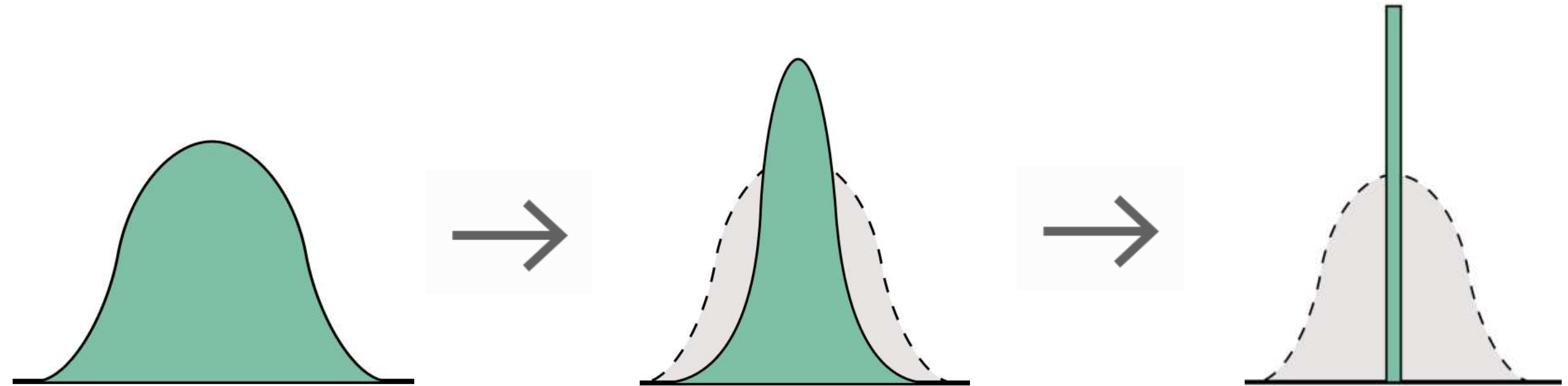
0:01 / 14:24



**Inside the Pantheon How Ancient Rome Built the 2,000 Year Old Impossible Dome**

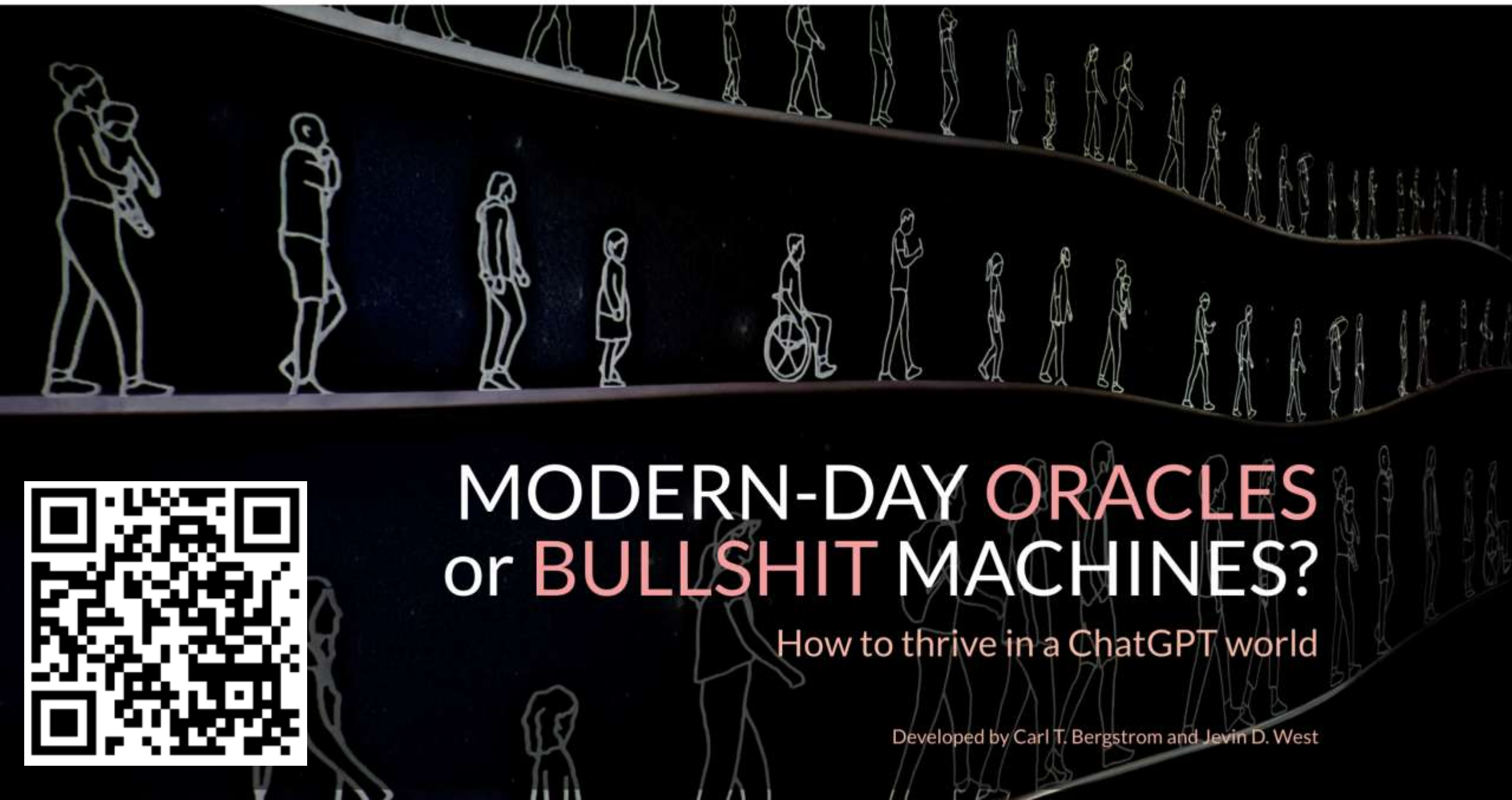
AI

# Epistemic diversity across language models mitigates *knowledge collapse*



# Public librarians as trusted AI intermediaries





# MODERN-DAY ORACLES or BULLSHIT MACHINES?

How to thrive in a ChatGPT world

Developed by Carl T. Bergstrom and Jevin D. West